KARNATAKA STATE AKKAMAHADEVI WOMEN'S UNIVERSITY, VIJAYAPURA

Research Proposal on Topic

"APPLICATIONS OF AI AND ML ALGORITHMS AS COUNTERMEASURES FOR INTERNET OF THINGS CYBER ATTACKS"

Submitted by

ASHA M

Research Scholar (Doctor of Philosophy)

(Reg.No: **PHCS2109**)

Research Guide

Dr. SHEETALRANI R KAWALE

Department of Computer Science

KSAWU, Vijayapura

DEPARTMENT OF COMPUTER SCIENCE

KARNATAKA STATE AKKAMAHADEVI WOMEN'S UNIVERSITY, VIJAYAPURA

2021-22

ABSTRACT

In recent years, the use of the Internet of Things (IoT) has increased exponentially, and cyber security concerns have increased along with it. On the cutting edge of cyber security is Artificial Intelligence (AI), which is used for the development of complex algorithms to protect networks and systems, including IoT systems. However, cyber-attackers have figured out how to exploit AI and have even begun to use adversarial AI in order to carry out cyber security attacks.

With the advances in information technology (IT) criminals are using cyberspace to commit numerous cybercrimes. Security and privacy of users have become significant concerns due to the involvement of the Internet of Things (IoT) devices in numerous applications. Cyber threats are growing at an explosive pace making the existing security and privacy measures inadequate.

Cyberspace has become an indispensable factor for all areas of the modern world. The world is becoming more and more dependent on the internet for everyday living. The increasing dependency on the internet has also widened the risks of malicious threats. On account of growing cyber security risks, cyber security has become the most pivotal element in the cyber world to battle against all cyber threats, attacks, and frauds. The expanding cyberspace is highly exposed to the intensifying possibility of being attacked by interminable cyber threats. In this research work we proposed to investigate the practices to reduce cyber security attacks for IoT using AI &ML approaches. An integrated IoT security framework that detects a variety of IoT attacks, as well as botnet attacks, based on various ML and DL models will be design. We also design a novel IDS/AI models to secure the IoT environment against cyber security attacks. Finally we will carry out the Validation of AI and ML approaches for providing security and privacy in IoT environments.

1. INTRODUCTION

The large-scale growth of the Internet of things (IoT) in recent years has contributed to a significant increase in fog computing, smart cities, and Industry 4.0, all of which execute the complex data processing of confidential information that must be protected against cyber security attacks. Cyber security attacks have increased rapidly in various domains, such as smart homes, healthcare, energy, agriculture, automation, and industrial processes [1]. As a result of their wide range of services, IoT device sensors generate large amounts of data that requires authentication, security, and privacy. Previously, traditional methods and frameworks were used to ensure the security of IoT. However, the application of different artificial intelligence (AI) methods for detecting cyber security attacks has gained in popularity over the years. The most important component used to detect cyber-attacks or malicious activities is the intrusion detection system (IDS). Artificial intelligence plays a vital role in detecting intrusions and widely considered as the better way in adapting and building IDS. In recent years, the use of the Internet of Things (IoT) has increased exponentially, and cyber security concerns have increased along with it. On the cutting edge of cyber security is Artificial Intelligence (AI), which is used for the development of complex algorithms to protect networks and systems, including IoT systems. However, cyber-attackers have figured out how to exploit AI and have even begun to use adversarial AI in order to carry out cyber security attacks. In order to dynamically protect systems from cyber threats, many cyber security experts are turning to Artificial Intelligence (AI). AI is most commonly used for intrusion detection in cyber security by analyzing traffic patterns and looking for an activity that is characteristic of an attack.

IoT comprises interconnected devices that are increasingly developed on a large scale, taking into account various characteristics through cloud and fog computing, where the processing of real-time applications can be enhanced [2]. Cyber-physical systems (CPSs) are integrated technologies such as healthcare IoT, industrial IoT (IIoT), smart cities IoT, AI and big data, that are part of the fourth industrial revolution (Industry 4.0), and are used for innovation in smart industries to promote data transmission between networks [3,4]. To overcome the security issues that threaten the IoT, several researchers have developed IDSs based on various AI approaches.

In this review, we intend to explore and analyze both ML and DL methods used for the detection of cyber security attacks in IoT devices and networks by formulating and addressing research questions, and filling the current research gaps in the area of IoT security by attempting to overcome limitations in the existing work. IDSs, particularly those designed using AI such as anomaly-based, are preferred for detecting cyber security attacks in mobile devices with integrity verification schemes for cloud storage in an endpoint environment [2, 29–31].

In this age, the cyberspace is growing faster as a primary source for a node to node information transfer with all its charms and challenges. The cyberspace serves as a significant source to access an infinite amount of information and resources over the globe. In 2017, the internet usage rate was 48% globally, later it increased to 81% for developing countries [1]. The broad spectrum of the cyberspace embraces the internet, users, the system resources, the technical skills of the participants and much more, not just the internet. The cyber-world also plays a significant role in causing limitless vulnerabilities to cyber threats and attacks. Cyber security is a set of different techniques, devices, and methods used to defend cyberspace against cyber-attacks and cyber threats [2]. In the modern world of computer and information technology, the cybercrimes are growing with faster steps as compared to the current cyber security system. The weak system

configuration, unskilled staff, and scanty amount of techniques are some factors that rise to vulnerabilities in a computer system to threats [3]. Because of the growing cyber threats, more headway needs to make when developing cyber security methods. The outdated and conventional cyber security methods have a substantial downside because these methods are ineffectual in dealing with unknown and polymorphic security attacks. There is a need for robust and advanced security methods that can learn from their experiences and detect the previous and new unknown attacks. Cyber threats are increasing in a significant way. It is becoming very challenging to cope with the speed of security threats and provide needful solutions to prevent them [4].

1. ARTIFICIAL INTELLIGENCE: The study and design of machines or computational methods that can perform tasks that normally require human intelligence.

2. GEO AI: An interdisciplinary field of geography and artificial intelligence.

3. MACHINE LEARNING: A sub field in AI that relies on statistical methods or numerical optimization techniques to derive models from data without explicitly programming every model parameter or computing step.

4. **DEEP LEARNING**: A special type of machine learning that leverages multiple layers of nonlinear processing units, or neurons, to learn representations from raw data to achieve the goal of automatic learning for completing various AI tasks.

Machine learning*:* One of the primarily used advanced methods for cybercrime detection is machine learning techniques. Machine learning techniques can be applied to address the limitations and constraints faced by conventional detection methods [5]. Researchers have addressed the advancements, limitations, and constraints of applying machine learning techniques for cyber attack detection and have provided a comparison of conventional methods

with machine learning techniques. Machine learning is a sub-field of artificial intelligence. ML techniques are built with the abilities to learn from experiences and data without being programmed explicitly [6]. Applications of ML techniques are expanding in different areas of life, such as education [7,8], medical [9–11], business and cyber security [12–14]. Machine learning techniques are playing their role on both sides of the net, i.e. attacker-side and defender-side. On the attacker side, ML techniques are employed to pass through the defense wall. In contrast, on defense side, ML techniques are applied to create prompt and robust defense strategies.

CYBER THREATS: Machine learning techniques are playing a vital role in fighting against cyber security threats and attacks such as intrusion detection system [15,16], malware detection [17], phishing detection [18,19], spam detection [20,21], and fraud detection [22] to name a few. We will focus on malware detection, intrusion detection system, and spam classification for this review. Malware is a set of instructions that are designed for malicious intent to disrupt the normal flow of computer activities. Malicious code runs on a targeted machine with the intent to harm and compromise the integrity, confidentiality and availability of computer resources and services [23]. Saad et al. in [24] discussed the main critical problems in applying machine learning techniques for malware detection. Saad et al. argued that machine learning techniques have the ability to detect polymorphic and new attacks. Machine learning techniques will lead to all other conventional detection methods in the future. The training methods for malware detections should be cost effective. The malware analysts should also be able to keep with the understanding of ML malware detection methods up to an expert level. Ambalavanan et al. in [25] described some of the strategies to detect cyber threats efficiently. One of the critical downsides of the security system is that the security reliability level of the computing resources

is generally determined by the ordinary user, who does not possess technical knowledge about security. Another threat to computer resources is a spam message. Spam messages are unwanted and solicited messages that consume a lot of network resources along with computer memory and speed. ML techniques are being employed to detect and classify a message as spam or ham. ML techniques have a significant contribution to detect spam messages on computer [26,27], SMS messages on mobile [28], spam tweets [29], or images/video [30,31]. An intrusion detection system (IDS) is a protection system to computer networks from any malign intrusions for scanning the network vulnerabilities. Signature-based, anomaly-based and hybrid-based are considered major classifications of an intrusion detection system for network analysis. ML techniques have a substantial contribution to detecting different types of intrusions on network and host computers. However, there are numerous areas such as detection of zero-day and new attacks are considered significant challenges for ML techniques [32].

COUNTERMEASURE: A countermeasure is an action, process, device, or system that can prevent, or mitigate the effects of, threats to a computer, server or network

CYBER SECURITY: Cyber security is a set of technologies and processes designed to protect computers, networks, programs and data from attack, damage, or unauthorized access [12]. In recent days, cyber security is undergoing massive shifts in technology and its operations in the context of computing, and data science (DS) is driving the change, where machine learning (ML), a core part of "Artificial Intelligence" (AI) can play a vital role to discover the insights from data. Machine learning can significantly change the cyber security landscape and data science is leading a new scientific paradigm. Craigen et al. defined "cyber security as a set of tools, practices, and guidelines that can be used to protect computer networks, software programs, and data from attack, damage, or unauthorized access" [24]. According to after good

et al. [12], “cyber security is a set of technologies and processes designed to protect computers, networks, programs and data from attacks and unauthorized access, alteration, or destruction”. Overall, cyber security concerns with the understanding of diverse cyber-attacks and devising corresponding defense strategies that preserve several properties defined as below [25, 26].

- **Confidentiality** is a property used to prevent the access and disclosure of information to unauthorized individuals, entities or systems.
- **Integrity** is a property used to prevent any modification or destruction of information in an unauthorized manner.
- **Availability** is a property used to ensure timely and reliable access of information assets and systems to an authorized entity. The term cyber security applies in a variety of contexts, from business to mobile computing, and can be divided into several common categories. These are - network security that mainly focuses on securing a computer network from cyber attackers or intruders; application security that takes into account keeping the software and the devices free of risks or cyber-threats; information security that mainly considers security and the privacy of relevant data; operational security that includes the processes of handling and protecting data assets. Typical cyber security systems are composed of network security systems and computer security systems containing a firewall, antivirus software, or an intrusion detection system [27].

Cyber-attacks and security risks: The risks typically associated with any attack, which considers three security factors, such as threats, i.e., who is attacking, vulnerabilities, i.e., the weaknesses they are attacking, and impacts, i.e., what the attack does [9]. A security incident is an act that threatens the confidentiality, integrity, or availability of information assets and systems. Several

types of cyber security incidents that may results in security risks on an organization's systems and networks or an individual [2].

These are:

• Unauthorized access that describes the act of accessing information to network, systems or data without authorization that results in a violation of a security policy [2];

• Malware known as malicious software, is any program or software that intentionally to cause damage to a computer, client, server, or computer network, e.g., botnets. Examples of different types of malware including computer viruses, worms, Trojan horses, adware, ransomware, spyware, malicious bots, etc. [3, 26]; Ransom malware, or ransomware, is an emerging form of malware that prevents users from accessing their systems or personal fles, or the devices, then demands an anonymous online payment in order to restore access.

• Denial-of-Service is an attack meant to shut down a machine or network, making it inaccessible to its intended users by fooding the target with trafc that triggers a crash. Te Denial-of-Service (DoS) attack typically uses one computer with an Internet connection, while distributed denial-of-service (DDoS) attack uses multiple computers and Internet connections to food the targeted resource [2]; • Phishing a type of social engineering, used for a broad range of malicious activities accomplished through human interactions, in which the fraudulent attempt takes part to obtain sensitive information such as banking and credit card details, login credentials, or personally identifiable information by disguising oneself as a trusted individual or entity via an electronic communication such as email, text, or instant message, etc. [26]; • Zero-day attack is considered as the term that is used to describe the threat of an unknown security vulnerability for which either the patch has not been released or the application developers were unaware [4, 28]. Beside these attacks mentioned above,

privilege escalation [29], password attack [30], insider threat [31], man-in-the-middle [32], advanced persistent threat [33], SQL injection attack [34], crypto jacking attack [35], web application attack [30] etc. are well known as security incidents in the Feld of cyber security. A data breach is another type of security incident, known as a data leak, which is involved in the unauthorized access of data by an individual, application, or service [5]. Thus, all data breaches are considered as security incidents; however, all the security incidents are not data breaches. Most data breaches occur in the banking industry involving the credit card numbers, personal information, followed by the healthcare sector and the public sector [36].

Problem Statement

The future Internet of Things (IoT) will have a deep economical, commercial and social impact on our lives. The participating nodes in IoT networks are usually resource constrained, which makes them luring targets for cyber-attacks. In this regard, extensive efforts have been made to address the security and privacy issues in IoT networks primarily through traditional cryptographic approaches. However, the unique characteristics of IoT nodes render the existing solutions insufficient to encompass the entire security spectrum of the IoT networks. This is, at least in part, because of the resource constraints, heterogeneity, massive real-time data generated by the IoT devices, and the extensively dynamic behavior of the networks. Therefore, Machine Learning (ML) and Deep Learning (DL) techniques, which are able to provide embedded intelligence in the IoT devices and networks, are leveraged to cope with different security problems.

In recent years, technology has advanced to the fourth industrial revolution (Industry 4.0), where the Internet of things (IoTs), fog computing, computer security, and cyber-attacks have evolved

exponentially on a large scale. The rapid development of IoT devices and networks in various forms generate enormous amounts of data which in turn demand careful authentication and security. Artificial intelligence (AI) is considered one of the most promising methods for addressing cyber security threats and providing security.

2. OBJECTIVES OF THE PROPOSED RESEARCH WORK

The objectives of the research work are

i. To investigate the practices to reduce cyber security attacks for IoT using AI & ML approaches
ii. To develop an integrated IoT security framework that detects a variety of IoT attacks, as well as botnet attacks, based on various ML and DL models.
iii. To improve the detection accuracy and concentrating on the availability of recent real-world datasets in order to detect new types of IoT threats
iv. To develop a novel IDS/AI models to secure the IoT environment against cyber security attacks.
v. Validation of AI and ML approaches for providing security and privacy in IoT environments.

3. LITERATURE REVIEW

A literature review includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. Literature review act as a guide for new researches and also helps in avoiding the rediscovery of known facts. The various related studies have been done in this field in India and the overseas and are discussed below:

i. ***Rachid Zagrouba and Reem Alhajri, Machine Learning based Attacks Detection and Countermeasures in IoT, International Journal of Communication Networks and Information Security (IJCNIS) Vol. 13, No. 2, August 2021***

In this paper, a low power consumption Machine Learning (ML) techniques have been proposed for detecting IoT botnet attacks using Random forest as ML-based detection method and describing IoT common attacks with its countermeasures. The experimental result shows higher accuracy. From the results, it is seen that IoT botnet detection is possible; achieving a higher accuracy rate as an experimental result indicates an accuracy rate of over 99.99% where the true positive rate is 1.000 and the false-negative rate is 0.000.

ii. ***NZ Jhanjhi, Mamoona Humayun, and Saleh N. Almuayqil, "Cyber Security and Privacy Issues in Industrial Internet of Things" tech science press, Computer Systems Science & Engineering, CSSE, 2021, vol.37, no.3, pp. 362-380, 2021.***

The focus of this paper is on investigating the possible cyber-attacks targeting 4 layers of IIoT that is one of the key pillars of Industry 4.0. Based on a detailed review of existing

literature, in this study, it has identified possible cyber threats, their consequences, and countermeasures. Further, authors have provided a comprehensive framework based on an analysis of cybersecurity and privacy challenges. The suggested framework provides for a deeper understanding of the current state of cyber security and sets out directions for future research and applications.

iii. ***Sundar Krishnan, Ashar Neyaz & Qingzhong Liu, "IoT Network Attack Detection using Supervised Machine Learning" International Journal of Artificial Intelligence and Expert Systems (IJAE), Volume(10) : Issue (2) : 2021 18 ISSN: 2180-124X,***

In this research, they perform various supervised feature selection methods and employ three classifiers on IoT network data. The classifiers predict with high accuracy if the network traffic against the IoT device was malicious or benign. A comparison has been made the feature selection methods to arrive at the best that can be used for network intrusion prediction.

iv. ***Qasem Abu Al-Haija and Saleh Zein-Sabatto, "An Efficient Deep-Learning-Based Detection and Classification System for Cyber-Attacks in IoT Communication Networks" Electronics 2020, 9, 2152; pp. 1-25, 2020, doi:10.3390/electronics9122152***

In this paper, it has been proposed a comprehensive development of a new intelligent and autonomous deep-learning-based detection and classification system for cyber-attacks in IoT communication networks that leverage the power of convolutional neural networks, IoT-IDCS-CNN (IoT based Intrusion Detection and Classification System using Convolutional Neural Network). To evaluate the developed system, Network Security Laboratory-Knowledge Discovery Databases (NSL-KDD) dataset, which includes all the

key attacks in IoT computing were used. The simulation results demonstrated a greater than 99.3% and 98.2% cyber-attack classification accuracy.

v. ***Jiyeon Kim, Minsun Shim , Seungah Hong , Yulim Shin and Eunjung Choi, "Intelligent Detection of IoT Botnets Using Machine Learning and Deep Learning" Appl. Sci. MDPI,PP. 1-22, 2020, 10, 7009; doi:10.3390/app10197009***

In this paper, a ML technique for IoT Attack detection is proposed. They focused on botnet attacks targeting various IoT devices and developed ML-based models for each type of device. We use the N-BaIoT dataset and developed a botnet detection model for each device using numerous ML models, including deep learning (DL) models. They then analyze the effective models with a high detection F1-score by carrying out multiclass classification, as well as binary classification, for each model.

vi. ***Ying Zhang , Peisong Li , and Xinheng Wang, "Intrusion Detection for IoT Based on Improved Genetic Algorithm and Deep Belief Network" IEEE Access, vol.7, pp. 31711-31722, 2019.***

This paper presents an intrusion detection model based on improved genetic algorithm (GA) and deep belief network (DBN). Facing different types of attacks, through multiple iterations of the GA, the optimal number of hidden layers and number of neurons in each layer are generated adaptively, so that the intrusion detection model based on the DBN achieves a high detection rate with a compact structure. Finally, the NSL-KDD dataset was used to simulate and evaluate the model and algorithms. The experimental results show that the improved intrusion detection model combined with DBN can effectively

improve the recognition rate of intrusion attacks and reduce the complexity of the neural network structure.

vii. ***S K. Niamathulla, Intrusion Detection with Machine Learning & Artificial Intelligence (ML&AI) Techniques to Reduce Cyberattacks (Network Traffic) (New Way to Improve Cybersecurity), International Journal of Science and Research (IJSR) ISSN: 2319-7064, Volume 8 Issue 11, pp. 85-101, November 2019***

In this paper, a Scalable Distributed Intrusion Detection System (DIDS) has been proposed using AI and ML techniques to improve cyber security and reduce cyber-attacks. The DIDS system gives the analyst (CSA/CTIA) a complete real-time accurate analysis of activities reports, it allows the analyst much more flexibility in discovering attack patterns. They propose to increase the bandwidth of network to improve the data rate flow of network traffic. For which it is easy to reduce the cyber-attacks on the network and save a lot of time and money.

viii. ***Jadel Alsamiri , Khalid Alsubhi, Internet of Things Cyber Attacks Detection using Machine Learning, International Journal of Advanced Computer Science and Applications, Vol. 10, No. 12,pp. 627-634, 2019***

In this study, the authors aim to contribute to the literature by evaluating various machine learning algorithms that can be used to quickly and effectively detect IoT network attacks. A new dataset, Bot-IoT, is used to evaluate various detection algorithms. In the implementation phase, seven different machine learning algorithms were used, and most

of them achieved high performance. New features were extracted from the Bot-IoT dataset during the implementation and compared with studies from the literature, and the new features gave better results.

ix. ***Furkan Yusuf YAVUZ1, Devrim ÜNA, Ensar GÜL, "Deep Learning for Detection of Routing Attacks in the Internet of Things", International Journal of Computational Intelligence Systems, Vol. 12 (2018) 39-58.***

In this paper, authors have proposed a deep-learning based machine learning method for detection of routing attacks for IoT. In this study, the Cooja IoT simulator has been utilized for generation of high-fidelity attack data, within IoT networks ranging from 10 to 1000 nodes. They propose a highly scalable, deep-learning based attack detection methodology for detection of IoT routing attacks which are decreased rank, hello-flood and version number modification attacks, with high accuracy and precision. Application of deep learning for cyber-security in IoT requires the availability of substantial IoT attack data and we believe that the IoT attack dataset produced in this work can be utilized for further research.

x. ***Miloud Bagaa , Tarik Taleb, Jorge Bernal Bernabe And Antonio Skarmeta, "A Machine Learning Security Framework for IoT Systems" Volume 4, pp. 1-12, IEEE Access, 2016.***

This paper presents a novel machine learning (ML) based security framework that automatically copes with the expanding security aspects related to IoT domain. This framework leverages both Software Defined Networking (SDN) and Network Function

Virtualization (NFV) enablers for mitigating different threats. Experiments results demonstrate the efficiency of the proposed scheme. In particular, the distribution of the attacks using the data mining approach is highly successful in detecting the attacks with high performance and low cost. Regarding our anomalybased intrusion detection system (IDS) for IoT, we have evaluated the experiment in a real Smart building scenario using one-class SVM. The detection accuracy of anomalies achieved 99.71%.

xi ***Hasan Alkahtani and Theyazn H. H. Aldhyani , Intrusion Detection System to Advance Internet of Things Infrastructure-Based Deep Learning Algorithms, Volume 2021, Article ID 5579851, 18 pages.***

In this paper robust framework system is proposed for detecting intrusion based on the IoT environment. The three advanced deep learning algorithms such as convolution neural network (CNN), a long short-term memory (LSTM), and a hybrid convolution neural network with the long short-term memory (CNN-LSTM) model were used to classify the intrusion. The particle swarm optimization method (PSO) was used to select relevant features from the network dataset. The experimental results showed that the proposed systems achieved accuracy as follows: CNN = 96.60%, LSTM= 99.82% and CNN-LSTM = 98.80%. The experimental results confirmed that the proposed framework can effectively detect real-world attacks and is capable of enhancing the security of the IoT environment.

xii ***Abdulkadir Bilen and Ahmet Bedri Özer, Cyber-attack method and perpetrator prediction using machine learning algorithms, PeerJ Comput. Sci., pp. 1-21, 2021, DOI 10.7717/peerj-cs.475***

In this paper, they proposed to analyze cybercrimes using eight machine learning methods and perpetrator prediction. The Support Vector Machine Linear was found out to be the most successful in the cyber-attack method, with an accuracy rate of 95.02%. The obtained results have revealed that the probability of cyber-attack decreases as the education and income level of victim increases. The proposed model facilitate the detection of cyber-attacks and make the fight against these attacks easier and more effective.

xiii ***John A. Bland, Mikel D. Petty, Tymaine S. Whitaker, Katia P. Maxwell, Walter Alan Cantrell, "Machine Learning Cyber-attack and Defense Strategies," Computers & Security, Vol. 92, pp.1-23, 2020***

In this work, cyber attacks were modeled using an extension of the well-known Petri net formalism. That formalism, designated Petri nets with players, strategies, and costs, models of the cyber attack and events during the attack as markings and transition firings in the net respectively. Given those validated nets, a reinforcement learning algorithm using an -Greedy policy was implemented and set to the task of learning which actions to take, These results demonstrate the potential of formally modeling cyber attacks and of applying reinforcement learning to improving cyber security

xiv ***Md Mamunur Rashi , Joarder Kamruzzaman, Mohammad Mehedi Hassan, Tasadduq Imam And Steven Gordon "Cyberattacks Detection In Iot-based Smart City Applications Using Machine Learning Techniques" Int. J. Environ. Res. Public Health 2020, 17, 9347, PP.1-21,2020***

In this paper, the feasibility of ensemble based learning with single model classifiers for identifying cyber attacks within the IoT-based smart city applications have been explored. The experiments with the most recent IoT attack database show that our ensemble approach, especially stacking, performs better than single models in identifying attacks from benign samples. The results indicating that stacking of classifiers can better detect cyber attacks in the smart city systems go beyond technical contributions and carry economic and social implications. Future research will provide further insights in this respect.

xv ***Farhan Ullah , Hamad Naeem1, Sohail Jabbar , Shehzad Khalid ,Muhammad Ahsan Latif , Fadi Al-turjman And Leonardo Mostarda "Cyber Security Threats Detection In Internet Of Things Using Deep Learning Approach" IEEE Access. Vol 7, Pp. 124379 – 124389, 2019***

In this paper, combined deep learning-based approach for the identification of pirated and malware files have been proposed. First, the Tens or Flow neural network is proposed to detect the pirated features of original software using software plagiarism. Secondly, a novel methodology based on convolution neural network and color image visualization to detect malware using IoT have been proposed. The experimental results show that the combined approach retrieve maximum classification results as compared to the state of the art techniques

xvi Mahmudul Hasan, Md. Milon Islam, Md Ishrak Islam Zarif, M.M.A. Hashem "Attack And Anomaly Detection In Iot Sensors In Iot Sites Using Machine Learning Approaches," Elsevier Internet Of Things 7, Pp. 1-14, 100059

In this paper, performances of several machine learning models have been compared to predict attacks and anomalies on the IoT systems accurately. The machine learning (ML) algorithms that have been used here are Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), and Artificial Neural Network (ANN). The system obtained 99.4% test accuracy for Decision Tree, Random Forest, and ANN. Though these techniques have the same accuracy, other metrics prove that Random Forest performs comparatively better.

xvii Vinayakumar R, Mamoun Alazab Soman Kp, Prabaharan Poornachandran , Ameer Al-nemrat , And Sitalakshmi Venkatraman, "Deep Learning Approach For Intelligentintrusion Detection System" IEEE Access 2018

In this paper, deep neural network (DNN), a type of deep learning model is explored to develop a flexible and effective IDS to detect and classify unforeseen and unpredictable cyber-attacks. A comprehensive evaluation of experiments of DNNs and other classical machine learning classifiers are shown on various publicly available benchmark malware datasets. Through a rigorous experimental testing it is confirmed that DNNs perform well in comparison to the classical machine learning classifiers. Finally, it has been proposed a highly scalable and hybrid DNNs framework called Scale-Hybrid-IDS-AlertNet (SHIA) which can be used in real time to effectively monitor the network traffic and host-level events to proactively alert possible cyber-attacks.

xviii Thomas Rincy N And Roopam Gupta, "Design And Development Of An Efficient Network Intrusion Detection System Using Machine Learning Techniques" Hindawi Wireless Communications And Mobile Computing, Pp. 1-35, Volume 2021

An efficient hybrid NID-Shield NIDS is proposed in this manuscript that classifies the UNSW-NB15 and NSL-KDD datasets according to the attack types and attack names. An effective hybrid feature subset selection method called CAPPER is applied as a feature subset selection that combines the CFS and Wrapper approaches for obtaining the reduced accurate and high merit feature subsets (iii) The reduced accurate and high merit datasets obtained from CAPPER are trained by the machine learning approaches and assessed by a 10-fold cross-validation method (iv) The hybrid NID-Shield network intrusion detection system shows overall good improvement results on the different approaches found in the existing literature studies. The hybrid NID-Shield NIDS with its various parameters is investigated with existing literature studies; it has been found that the hybrid NID-Shield NIDS is the most efficient of all approaches found in the existing literature stud.

ixx Madhavi Dhingra, S C Jain, Rakesh Singh Jadon "Malicious Intrusion Detection Using Machine Learning Schemes" International Journal Of Engineering And Advanced Technology (Ijeat) Volume-8 Issue-6, August 2019

In this paper a ML is used for malicious intrusion detection.The proposed research work has implemented the Intelligent machine learning classification algorithms on UNSW-NB 15 data set after transforming the dataset by using Correlation feature selection algorithm. The results indicate that the classifier Multilayer Perceptron is giving best

True positive rate and ROC Area, whereas Bagging and RepTree are giving remarkable results on the reduced dataset

xx ***Abel Yeboah-ofori, Classification Of Malware Attacks Using Machine Learning In Decision Tree, International Journal Of Security (IJS), Volume (11) : Issue (2), Pp. 10-24, 2020***

This paper explores the viability of using machine learning methods to predict malware attacks and build a classifier to automatically detect and label an event as "Has Detection or No Detection". The purpose is to predict the probability of malware penetration and the extent of manipulation on the network nodes for cyber threat intelligence. To demonstrate the applicability of our work, we use a decision tree (DT) algorithms to learn dataset for evaluation. The dataset was from Microsoft Malware threat prediction website Kaggle. We identify probably cyberattacks on smart grid, use attack scenarios to determine penetrations and manipulations. The results show that ML methods can be applied in smart grid cyber supply chain environment to detect cyberattacks and predict future trends

4. PROPOSED SYSTEM AND METHODOLOGY

i. Proposed system:

We propose to investigate the practices to reduce cyber security attacks for IoT using AI and ML approaches. Also, we propose to develop integrated IoT security framework and novel IDS/Ai models. We will use recent real-world datasets in order to detect new types of IoT threats. Then we develop a novel IDS/AI models to secure the IoT environment against cyber security attacks. Finally, we carry out the validation of AI and ML approaches for providing security and privacy in IoT environments.

Threats to validity: we use anyone of the six machine learning models in cyber security and target cyber threats including intrusion detection, spam detection, and malware detection. The performance evaluation in terms of accuracy, recall, or precision is carried out.

ii. Methodology:

The overall framework is a combination of several independent processes. Fig. 1 depicts the overall framework of the system. The first process of this framework is the dataset collection and dataset observation. In this process, the dataset will be collected and observed meticulously to find out the types of data. Besides, data preprocessing would be implemented on the dataset. Data preprocessing consists of cleaning of data, visualization of data, feature engineering and vectorization steps. These steps convert the data into feature vectors. These feature vectors are then split into 80–20 ratio into training and testing set. The training set is used in Learning Algorithm, and a final model is developed using an optimization technique. Different classifiers that will be used in this work employee different optimization techniques. The final model was evaluated against the testing set using different evaluation metrics.

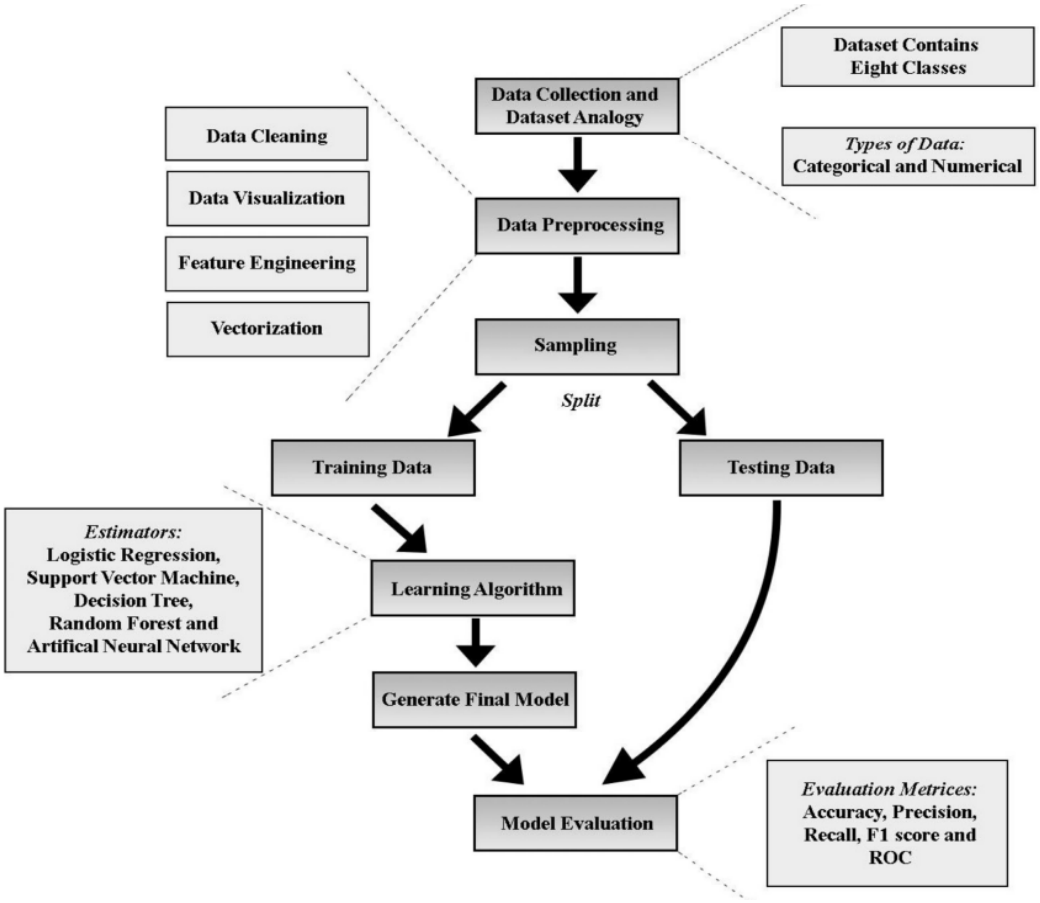

Figure 1: Overall framework for attack and anomaly detection in IoT.

iii. Cyber security Datasets :

In this research, the variables of interest are concerned with the intrusions/attacks on data records in IoT computing environments. Two global datasets of IoT attacks can be investigated, including the KDD'99 dataset and the NSL-KDD dataset. Indeed, KDD'99 has been developed by the Defense Advanced Research Projects Agency(DARPA) intrusion detection evaluation program to build a network IDS that is capable of differentiating between "bad" and "good" connections [27]

Importing the NSL-KDD dataset: In this stage, the collected dataset was imported/read using MATLAB 2019b (by Math Works, Inc.) in a tabulated format instead of raw data in the original dataset text files.

One of the most widely used datasets for intrusion detection is the Knowledge Discovery and Dissemination (KDD) 1999 dataset [57]. This dataset was created for the KDD Cup challenge in 1999 and is composed of more than 4 million network traffic records. However, the dataset does not include raw network traffic data. Instead, the raw cap data has been preprocessed into 41 features based on basic type, content type, and traffic type features. The dataset contains 22 different kinds of attacks that fall into four families: denial of service (DoS), unauthorized access from a remote machine (R2L), unauthorized access to local super-user privileges (U2R), and probing. An analysis by [58], however, found there were a number of problems with the dataset. One major source of this was the synthetic nature of the network and attack data, dropped data due to overflow, and vague attack definitions. Furthermore, there were a large number of redundant records that biased the dataset. Because of these flaws, [58] proposed a new dataset called NSL-KDD, which is another frequently used dataset for network intrusion detection.

5. EXPECTED OUTCOME

The proposed system will be designed and implemented using the latest MATLAB tool. The integrated IoT security framework and novel IDS/Ai models are developed using MATLAB. The results of this simulation will demonstrate the effectiveness novel models for securing the IoT environment against cyber security attacks. Also the results will show the validation of AI and ML approaches for providing security and privacy in IoT environments. We will use anyone of the six machine learning models in cyber security and target cyber threats including intrusion

detection, spam detection, and malware detection. The performance evaluation of machine learning models will be carried out in terms of accuracy, f-measure, recall, or precision will be carried out.

REFERENCES

1. Singh, S.; Sheng, Q.Z.; Benkhelifa, E.; Lloret, J. Guest Editorial: Energy Management, Protocols, and Security for the NextGeneration Networks and Internet of Things. IEEE Trans. Ind. Inform. 2020, 16, 3515–3520. [CrossRef]
2. Almiani, M.; AbuGhazleh, A.; Al-Rahayfeh, A.; Atiewi, S.; Razaque, A. Deep recurrent neural network for IoT intrusion detection system. Simul. Model. Pract. Theory 2020, 101, 102031. [CrossRef]
3. Hong, Z.; Hong, M.; Wang, N.; Ma, Y.; Zhou, X.; Wang, W. A wearable-based posture recognition system with AI-assisted approach for healthcare IoT. Futur. Gener. Comput. Syst. 2022, 127, 286–296. [CrossRef]
4. Adil, M.; Khan, M.K. Emerging IoT Applications in Sustainable Smart Cities for COVID-19: Network Security and Data Preservation Challenges with Future Directions. Sustain. Cities Soc. 2021, 75, 103311. [CrossRef] [PubMed]
5. Kurte, R.; Salcic, Z.; Wang, K.I.K. A Distributed Service Framework for the Internet of Things. IEEE Trans. Ind. Inform. 2020, 16,4166–4176. [CrossRef]
6. Zeng, P.; Pan, B.; Choo, K.K.R.; Liu, H. MMDA: Multidimensional and multidirectional data aggregation for edge computingenhanced IoT. J. Syst. Archit. 2020, 106, 101713. [CrossRef]
7. Diro, A.A.; Chilamkurti, N. Distributed attack detection scheme using deep learning approach for Internet of Things. Futur. Gener. Comput. Syst. 2018, 82, 761–768. [CrossRef]

8. Farivar, F.; Haghighi, M.S.; Jolfaei, A.; Alazab, M. Artificial Intelligence for Detection, Estimation, and Compensation of Malicious Attacks in Nonlinear Cyber-Physical Systems and Industrial IoT. IEEE Trans. Ind. Inform. 2020, 16, 2716–2725. [CrossRef]

9. Gupta, M.; Abdelsalam, M.; Khorsandroo, S.; Mittal, S. Security and Privacy in Smart Farming: Challenges and Opportunities. IEEE Access 2020, 8, 34564–34584. [CrossRef]

10. Al-Haija, Q.A.; Zein-Sabatto, S. An efficient deep-learning-based detection and classification system for cyber-attacks in iot communication networks. Electronics 2020, 9, 2152. [CrossRef]

11. Zhang, T.; Zhao, Y.; Jia, W.; Chen, M.Y. Collaborative algorithms that combine AI with IoT towards monitoring and control system. Futur. Gener. Comput. Syst. 2021, 125, 677–686. [CrossRef]

12. Li, B.; Feng, Y.; Xiong, Z.; Yang, W.; Liu, G. Research on AI security enhanced encryption algorithm of autonomous IoT systems. Inf. Sci. 2021, 575, 379–398. [CrossRef]

13. Karale, A. The Challenges of IoT Addressing Security, Ethics, Privacy, and Laws. Internet Things 2021, 15, 100420. [CrossRef]

14. Obaidat, M.A.; Obeidat, S.; Holst, J.; Hayajneh, A.A.; Brown, J. A comprehensive and systematic survey on the internet of things: Security and privacy challenges, security frameworks, enabling technologies, threats, vulnerabilities and countermeasures. Computers 2020, 9, 44. [CrossRef]

15. Li, J.; Kuang, X.; Lin, S.; Ma, X.; Tang, Y. Privacy preservation for machine learning training and classification based on homomorphic encryption schemes. Inf. Sci. 2020, 526, 166–179. [CrossRef]

16. Sarica, A.K.; Angin, P. Explainable security in SDN-based IoT networks. Sensors 2020, 20, 7326. [CrossRef] [PubMed]

17. Aleem, S.; Capretz, L.F.; Ahmed, F. Security Issues in Data Warehouse. arXiv 2015, arXiv:1507.05644.

18. Wu, M.; Song, Z.; Moon, Y.B. Detecting cyber-physical attacks in CyberManufacturing systems with machine learning methods. J. Intell. Manuf. 2019, 30, 1111–1123. [CrossRef]

19. Patil, R.; Dudeja, H.; Modi, C. Designing in-VM-assisted lightweight agent-based malware detection framework for securing virtual machines in cloud computing. Int. J. Inf. Secur. 2020, 19, 147–162. [CrossRef]

20. Dang, T.K.; Pham, C.D.M.; Nguyen, T.L.P. A pragmatic elliptic curve cryptography-based extension for energy-efficient device-todevice communications in smart cities. Sustain. Cities Soc. 2020, 56, 102097. [CrossRef]

21. Moustafa, N. A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets. Sustain. Cities Soc. 2021, 72, 102994. [CrossRef]

22. Atul, D.J.; Kamalraj, R.; Ramesh, G.; Sakthidasan Sankaran, K.; Sharma, S.; Khasim, S. A machine learning based IoT for providing an intrusion detection system for security. Microprocess. Microsyst. 2021, 82, 103741. [CrossRef]

23. Ghosh, A.; Chakraborty, D.; Law, A. Artificial intelligence in Internet of things. CAAI Trans. Intell. Technol. 2018, 3, 208–218. [CrossRef]

24. Bland, J.A.; Petty, M.D.; Whitaker, T.S.; Maxwell, K.P.; Cantrell, W.A. Machine Learning Cyberattack and Defense Strategies. Comput. Secur. 2020, 92, 101738. [CrossRef]

25. Rathore, S.; Park, J.H. Semi-supervised learning based distributed attack detection framework for IoT. Appl. Soft Comput. J. 2018, 72, 79–89. [CrossRef]

26. Kasongo, S.M.; Sun, Y. A deep learning method with wrapper based feature extraction for wireless intrusion detection system. Comput. Secur. 2020, 92, 101752. [CrossRef]

27. Chmiel, M.; Korona, M.; Kozioł, F.; Szczypiorski, K.; Rawski, M. Discussion on iot security recommendations against the state-of-the-art solutions. Electronics 2021, 10, 1814. [CrossRef]

28. Aldhaheri, S.; Alghazzawi, D.; Cheng, L.; Barnawi, A.; Alzahrani, B.A. Artificial Immune Systems approaches to secure the internet of things: A systematic review of the literature and recommendations for future research. J. Netw. Comput. Appl. 2020, 157, 102537. [CrossRef]

29. Quintal, K.; Malton, A.; Walenstein, A. Biometric Signatures for Continuous Authentication. Digit. Object Identifier 2019, 23, 18–28. [CrossRef]

30. Lu, X.; Pan, Z.; Xian, H. An integrity verification scheme of cloud storage for internet-of-things mobile terminal devices. Comput. Secur. 2020, 92, 101686. [CrossRef]

31. Kim, S.; Hwang, C.; Lee, T. Anomaly based unknown intrusion detection in endpoint environments. Electronics 2020, 9, 1022. [CrossRef]

32. Choo, K.K.R.; Yan, Z.; Meng, W. Editorial: Blockchain in Industrial IoT Applications: Security and Privacy Advances, Challenges, and Opportunities. IEEE Trans. Ind. Inform. 2020, 16, 4119–4121. [CrossRef]

33. Almusaylim, Z.A.; Alhumam, A.; Jhanjhi, N.Z. Proposing a Secure RPL based Internet of Things Routing Protocol: A Review. Ad Hoc Netw. 2020, 101, 102096. [CrossRef]

34. Viejo, A.; Sánchez, D. Secure monitoring in IoT-based services via fog orchestration. Futur. Gener. Comput. Syst. 2020, 107, 443–457. [CrossRef]

35. Farahmand, F.; Navathe, S.B.; Enslow, P.H.; Sharp, G.P. Managing vulnerabilities of information systems to security incidents. In Proceedings of the 5th International Conference on Electronic Commerce, Pittsburgh, PA, USA, 30 September–3 October 2003; pp. 348–354.

36. Szor, P. The Art of Computer Virus Research and Defense: ART COMP VIRUS RES DEFENSE _p1; Pearson Education: London, UK, 2005.

37. Firdausi, I.; Erwin, A.; Nugroho, A.S. Analysis of machine learning techniques used in behavior-based malware detection. In Proceedings of the 2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies, Jakarta, Indonesia, 2–3 December 2010; pp. 201–203.

38. Shaukat, K.; Nawaz, I.; Zaheer, S. Students Performance: A Data Mining Perspective; LAP Lambert Academic Publishing: Saarbrücken, Germany, 2017.

39. Shaukat, K.; Nawaz, I.; Aslam, S.; Zaheer, S.; Shaukat, U. Student's performance in the context of data mining. In Proceedings of the 2016 19th International Multi-Topic Conference (INMIC), Islamabad, Pakistan, 5–6 December 2016; pp. 1–8.

40. Shaukat, K.; Masood, N.; Mehreen, S.; Azmeen, U. Dengue fever prediction: A data mining problem. J. Data Min. Genom. Proteom. 2015, 2015, doi:10.4172/2153-0602.1000181.

41. Jusas, V.; Samuvel, S.G. Classification of motor imagery using combination of feature extraction and reduction methods for brain-computer interface. Inf. Technol. Control 2019, 48, 225–234.

42. Uktveris, T.; Jusas, V. Comparison of feature extraction methods for EEG BCI classification. In Proceedings of the International Conference on Information and Software Technologies, Vilnius, Lithuania, 10–12 October 2015; pp. 81–92.

43. . Shaukat, K.; Rubab, A.; Shehzadi, I.; Iqbal, R. A Socio-Technological analysis of Cyber Crime and Cyber Security in Pakistan. Transylv. Rev. 2017, 1, 84.

44. Canhoto, A.I.; Clear, F. Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. Bus. Horiz. 2019, 63, 183–193.

45. Maqsood, H.; Mehmood, I.; Maqsood, M.; Yasir, M.; Afzal, S.; Aadil, F.; Selim, M.M.; Muhammad, K. A local and global event sentiment based efficient stock exchange forecasting using deep learning. Int. J. Inf. Manag. 2020, 50, 432–451.

46. Dey, S.; Ye, Q.; Sampalli, S. A machine learning based intrusion detection scheme for data fusion in mobile clouds involving heterogeneous client networks. Inf. Fusion 2019, 49, 205–215.

47. Geluvaraj, B.; Satwik, P.; Kumar, T.A. The future of cybersecurity: Major role of artificial intelligence, machine learning, and deep learning in cyberspace. In International Conference on Computer Networks and Communication Technologies; Springer: Berlin/Heidelberg, Germany, 2018; pp. 739–747.

48. Jain, P. Machine Learning Versus Deep Learning for Malware Detection. Master's Thesis, San Jose State University, San Jose, CA, USA, 2019.

49. Rao, R.S.; Pais, A.R. Detection of phishing websites using an efficient feature-based machine learning framework. Neural Comput. Appl. 2019, 3851–3873.

50. Alauthman, M.; Almomani, A.; Alweshah, M.; Omoushd, W.; Alieyane, K. Machine Learning for phishing Detection and Mitigation. Mach. Learn. Comput. Cyber Secur. Princ. Algorithmsand Pract. 2019, 26, 48–74.

51. Alurkar, A.A.; Ranade, S.B.; Joshi, S.V.; Ranade, S.S.; Shinde, G.R.; Sonewar, P.A.; Mahalle, P.N. A Comparative Analysis and Discussion of Email Spam Classification Methods Using Machine Learning Techniques. In Applied Machine Learning for Smart Data Analysis; CRC Press:Boca Raton, FL, USA, 2019; p. 185.

52. Dada, E.G.; Bassi, J.S.; Chiroma, H.; Adetunmbi, A.O.; Ajibuwa, O.E. Machine learning for email spam filtering: Review, approaches and open research problems. Heliyon 2019, 5, e01802.

53. Shukur, H.A.; Kurnaz, S. Credit Card Fraud Detection Using Machine Learning Methodology. Int. J. Comput. Sci. Mob. Comput. 2019, 8, 257–260. Energies 2020, 13, 2509 24 of 30

54. Afek, Y.; Bremler-Barr, A.; Feibish, S.L. Zero-day signature extraction for high-volume attacks. IEEE/ACM Trans. Netw. 2019, 27, 691–706.

55. Saad, S.; Briguglio, W.; Elmiligi, H. The Curious Case of Machine Learning In Malware Detection. arXiv 2019, arXiv:1905.07573

56. Ambalavanan, V. Cyber Threats Detection and Mitigation Using Machine Learning. In Handbook of Research on Machine and Deep Learning Applications for Cyber Security; IGI Global: Hershey, PA, USA, 2020; pp. 132– 149.

57. Shah, N.F.; Kumar, P. A comparative analysis of various spam classifications. In Progress in Intelligent Computing Techniques: Theory, Practice, and Applications; Springer: Berlin/Heidelberg, Germany, 2018; pp. 265–271.

58. . Chandrasekar, C.; Priyatharsini, P. Classification techniques using spam filtering email. Int. J. Adv. Res. Comput. Sci. 2018, 9, 402.

59. Shafi'I, M.A.; Latiff, M.S.A.; Chiroma, H.; Osho, O.; Abdul-Salaam, G.; Abubakar, A.I.; Herawan, T. A review on mobile SMS spam filtering techniques. IEEE Access 2017, 5, 15650–15666.

60. Chen, C.; Zhang, J.; Xie, Y.; Xiang, Y.; Zhou, W.; Hassan, M.M.; AlElaiwi, A.; Alrubaian, M. A performance evaluation of machine learning-based streaming spam tweets detection. IEEE Trans. Comput. Soc. Syst. 2015, 2, 65–76.

61. Biggio, B.; Fumera, G.; Pillai, I.; Roli, F. A survey and experimental evaluation of image spam filtering techniques. Pattern Recognit. Lett. 2011,

62. , 1436–1446. 31. Kumar, A.D.; KP, S. DeepImageSpam: Deep Learning based Image Spam Detection. arXiv 2018, arXiv:1810.03977.

63. . Jusas, V.; Japertas, S.; Baksys, T.; Bhandari, S. Logical filter approach for early stage cyber-attack detection. Comput. Sci. Inf. Syst. 2019, 16, 491–514.

64. Xin, Y.; Kong, L.; Liu, Z.; Chen, Y.; Li, Y.; Zhu, H.; Gao, M.; Hou, H.; Wang, C. Machine learning and deep learning methods for cybersecurity. IEEE Access 2018, 6, 35365–35381.

65. Gandotra, E.; Bansal, D.; Sofat, S. Malware analysis and classification: A survey. J. Inf. Secur. 2014, 5, 56.